PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [7] :	A2	(11) International Publication Number:	**WO 00/25827**
A61K 48/00		(43) International Publication Date:	11 May 2000 (11.05.00)

(21) International Application Number: PCT/EP99/07874

(22) International Filing Date: 18 October 1999 (18.10.99)

(30) Priority Data:
MI98A002330 30 October 1998 (30.10.98) IT

(71) Applicant *(for all designated States except US)*: MENARINI RICERCHE S.P.A. [IT/IT]; Via Tito Speri, 10, I-00040 Pomezia (IT).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: PARENTE, Dino [IT/IT]; (IT). DI MASSIMO, Anna, Maria [IT/IT]; (IT). DE SANTIS, Rita [IT/IT]; Via Rismondo, I-50131 Firenze (IT).

(74) Agent: MINOJA, Fabrizio; Bianchetti Bracco Minoja Srl, Via Rossini, 8, I-20122 Milano (IT).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PHARMACEUTICAL COMPOSITION, CONTAINING FRAGMENTS OF AN ANTIGENIC PROTEIN ENCODING DNA ENDOWED WITH ANTI-TUMOR EFFECT

(57) Abstract

Provided herein is a pharmaceutical composition containing one or more DNA molecules encoding fragments of a protein overexpressed in tumor cells, in order to induce an anti-tumor Ag-specific immune response, in association with suitable excipients and adjuvants.

PHARMACEUTICAL COMPOSITION, CONTAINING FRAGMENTS OF AN ANTIGENIC
PROTEIN ENCODING DNA ENDOWED WITH ANTI-TUMOR EFFECT.

Field of the invention

5   The invention relates to a pool of DNA plasmid constructs
containing the sequences of human MUC-1 encoding fragments and
to a pool of DNA plasmids in which the fragments themselves are
preceded by the sequence encoding a protein consisting of human
ubiquitin fused to a bacterial LacI fragment. The invention
10   further relates to their use in the preparation of
pharmaceutical compositions for use as DNA anti-tumor vaccines.

Background art

The invention provides an anti-tumor therapy based on the
induction or activation of the immune response able to bring
15   about tumor rejection. The validity of such an idea is
demonstrated from the first clinical results; for example,
patients treated with a viral vaccine containing the
Carcinoembryonic Antigen (CEA) encoding sequences demonstrated
immune system activation against this antigen (Tsang KY et al.
20   J. Natl. Cancer. Inst. 87: 982, 1995).

The activation of an immune anti-tumor response is
achievable through four different approaches:

a) Ex vivo engineering of patient tumor cells in order to
make them more immunogenic and suitable as a vaccine;

25   b) Ex vivo engineering of patient immune cells in order to
pre-activate an in vitro immune response.

c) Inoculation of naked or liposome capsulated or viral
particle integrated (retrovirus, vaccinia virus, adenovirus,
etc.) DNA encoding tumor associated antigens;

30   d) Treatment with recombinant or synthetic soluble tumor
antigens conjugated or mixed with adjuvants.

The first two approaches consist of the engineering of
every single patient cell and are limited in that they are
necessarily patient-specific, while the latter two are aimed to

obtain products comparable to a traditional drug.

The new vaccination methods reflect the development of new
technologies. The recent indications coming from the
experimentation on DNA naked vaccines that induce either a
5   persistent antibody or a cell immune response, make the
traditional protein subunit vaccines constituted of certain
specific peptides, inducing a lymphocyte population, obsolete.
Intramuscularly or intradermically injected proteins, encoded by
naked DNA, induce a cytotoxic-specific response as well as a
10   helper response. This powerful combination is extremely
effective but the underling mechanism is not completely
clarified yet. Muscle cells express class I MHC antigens at low
levels only, and do not apparently express class II antigens or
co-stimulatory molecules. Consequently, transfected muscle cells
15   are unlikely to play an important role in the onset of the
immune response per se. Recent data show that Antigen Presenting
Cells (APC), such as macrophages or dendritic cells, play a
fundamental role in capturing the myocyte released antigen and
in the subsequent processing and presenting of the respective
20   peptides in the context of the class I and II molecules, thus
inducing a CD8+ cell activation with cytotoxic activity as well
as activation of the CD4+ cells co-operating with B lymphocytes
in eliciting the antibody response (Corr M et al J. Exp. Med.
184:1555, 1996) (Tighe, H. et al. Immunology Today 19:89, 1998).

25   Furthermore, the use of cytokines is known to improve the
therapeutic effect deriving from immunization with DNA.
Cytokines can be administered in the form of exogenous proteins
as reported in Irvine et al., J. Immunol. 156: 238, 1996. An
alternative approach is represented by the contemporaneous
30   inoculation of both the tumor antigen or the desired cytokine
encoding plasmids, thus allowing the cytokine to be produced in
situ (Kim JJ et al. Immunol 158: 816, 1997).

The active immunization approach of the present invention
is based on the use of DNA vectors as vaccines against the MUC-1

human   antigen   or   Polymorphic   Epithelial   Mucin   (PEM),
overexpressed in tumor cells. MUC-1 is an epithelial luminal
surface glycoprotein (Patton S. et al. BBA 1241:407, 1995). In
the cell transformation process this glycoprotein loses the
apical localization and its expression level rises dramatically.
The protein function consists of protecting the luminal
surfaces, for example in the mammal gland, ovary, endometrium,
colon, stomach, pancreas, bladder, kidney, etc. A glycosylation
defect is reported that makes tumor cell associated MUC-1
antigenically different from normal cell associated MUC-1. This
phenomenon causes tumor MUC-1 to expose the antigen epitopes
that are normally masked by the sugar moieties in the normal
cell expressed MUC-1. This characteristic makes tumor MUC-1
particularly interesting in an induction of a tumor specific
antibody response (Apostolopoulos V. et al. Crit. Rev. Immunol.
14:293, 1994).

As an objective, the vaccination is aimed at inducing
immune responses against tumor cells expressing MUC1 at high
levels, preserving at the same time the low expressing normal
epithelia. The DNA vaccination relies upon the entrance of a
gene or portions thereof inside the body cells followed by
transcription and translation of the inserted sequence and thus
the intracellular synthesis of the corresponding polypeptide. An
important advantage of this system is that the neo-synthesized
protein is naturally processed inside the cell and the produced
peptides are associated with the Major Histocompatibility
Complex class I molecules (MHC-I). The MHC/peptide complexes are
therefore naturally exported to the cell surface where they can
be recognized by the immune system CD8+ cytotoxic cells. Only
the polypeptides synthesized inside the cell are then processed
and presented in association with the MHC class I molecules,
thus making it the only mechanism to stimulate, a specific
cytotoxic response. Vaccination systems based on protein or
peptide administration are usually more effective in stimulating

4

the antibody immune response which, to date, has been shown to
be ineffective in rejecting tumor cells. Current gene therapy
techniques rely upon DNA packaging in recombinant viral vectors
(retrovirus and adenovirus). The naked DNA administration is
much more advantageous in terms of effectiveness and safety
compared to viral vector therapies (*Kumar V and Sercarz E.
Nature Med. 2: 857, 1996; McDonnel WM et al., New England J. of
Med. 334: 42, 1996*). In fact naked DNA is unable either to
duplicate or integrate in the host tissue DNA and does not
induce the immune response to viral proteins.

The use of the ubiquitin to enhance the neo-synthesized
protein processing and thus cytotoxic lymphocyte induction was
recently reported (*Rodriguez F. et al., J. Virology 71: 8497,
1997*). The use of ubiquitin in order to generate proteins with
an N-terminal amino acid, making them unstable and thus prone to
enhanced degradation, had been previously reported (*Bechmair A.
et al., SCIENCE 234: 179, 1986*). The higher instability of these
proteins was subsequently related to enhanced intracellular
processing and presentation of model proteins by MHC-1 (*Grant E
P et al., J. Immunol. 155: 3750, 1995*) (*Wu Y and Kipps T.J., J.
Immunol. 159: 6037, 1997*).

The use of single constructs containing partial antigen
encoding DNA fragments (influenza virus nucleoprotein), having a
higher antigenic presentation efficiency compared to the
analogues with the whole antigenic sequence, in DNA vaccination
was reported (*Anton L. C. et al., J. Immunol. 158: 2535, 1997*).
Furthermore the processing of intracellular proteins and
presentation of the respective peptides by MHC class I proteins
in physiologic conditions, underlie the mechanism of
immunological surveillance. For a given protein and a specific
MHC context, there are peptide fragments termed dominants (i. e.
prevailing on subdominants or cryptics), which are unable to
generate any immune response because they are recognized as
"self". It has now been outlined, according to an aspect of the

present invention, that an approach aimed at supporting the non-dominant epitope presentation by the administration of a mix of antigen protein fragments is able to elicit a surprising cytotoxic immune response.

5   **Description of the invention**

It has now been found that DNA molecules, encoding fragments of a protein overexpressed in tumor cells, can be conveniently used to induce an antigen-specific anti-tumor immune response.

10   The invention relates particularly to a pharmaceutical composition containing one or more DNA encoding Mucin (MUC-1) protein fragments.

The DNA used in the present invention can be plasmid or viral DNA, preferably plasmid DNA obtained employing the pMRS30 15 expression vector described in fig. 13.

The compositions according to the invention contain preferably at least two DNA fragments of the Mucin (MUC-1) or of another protein overexpressed in tumor cells.

The compositions according to the invention contain 20 preferably at least four fragments, each ranging from 200 to about 700 nucleotides, each sequence being juxtaposed and possibly partially overlapping, from about 50 to about 150 nucleotides, at the 3' and/or 5' end of the adjacent one.

The DNA fragments according to the invention can be 25 possibly preceded at the 5' end by a ubiquitin encoding DNA sequence and possibly also by a LacI portion of Escherichia coli.

The invention relates also to new DNA fragments and to the use of Mucin-1 fragments defined above in the medicine and anti-30 tumor vaccine preparation.

Description of the figures

Fig. 1

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS166 expression

vector. This DNA includes the sequence corresponding to nucleotides 136-339 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by the two translation stop codons, TGA and TAA. The encoded polypeptide thus includes

5   a Metionin followed by the amino acids encoded by the 136-339 fragment of the EMBL sequence J05581.

Fig. 2

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression

10   vector to give the pMRS169 expression vector. This DNA includes the sequence corresponding to nucleotides 205-720 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes a Metionin followed by the

15   amino acids encoded by the 205-720 fragment of the EMBL sequence J05581.

Fig. 3

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression

20   vector to give the pMRS168 expression vector. This DNA includes the sequence corresponding to nucleotides 631-1275 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes a Metionin followed by the

25   amino acids encoded by the 631-1275 fragment of the EMBL sequence J05581.

Fig. 4

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression

30   vector to give the pMRS167 expression vector. This DNA includes the sequence corresponding to nucleotides 1222-1497 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes a Metionin followed by the

amino acids encoded by the 1222-1497 fragment of the EMBL sequence J05581.

Fig. 5

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS175 expression vector. This DNA includes the sequence corresponding to nucleotides 136-1497 of the EMBL sequence J05581, preceded by the translation start codon, ATG and followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes a Metionin followed by the amino acids encoded by the 136-1497 fragment of the EMBL sequence J05581.

Fig. 6

Nucleotide DNA sequence (with the respective amino acid sequence) termed UBILacI. The encoded polypeptide includes the Ubiquitin sequence fused to a partial sequence of the bacterial protein beta-galactosidase, as described in *Chau V. et al. Science 243: 1576, 1989*.

Fig. 7

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the expression vector pMRS30 to give the pMRS171 expression vector. This DNA includes the sequence termed UBILacI (see fig. 6) fused to the sequence corresponding to nucleotides 136-339 of the EMBL sequence J05581 followed by two translation stop codons, TGA and TAA. The coded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 136-339 of the EMBL sequence J05581.

Fig. 8

Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS174 expression vector. This DNA includes the sequence termed UBILacI (see fig. 6) fused to the sequence partially corresponding to nucleotides 205-720 of the EMBL

sequence J05581 followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 205-720 of the EMBL sequence

5   J05581.

       Fig. 9

       Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS173 expression vector. This DNA includes

10   the sequence termed UBILacI (see fig. 6) fused to the sequence partially corresponding to nucleotides 631-1275 of the EMBL sequence J05581 followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the

15   amino acids encoded by the fragment 631-1275 of the EMBL sequence J05581.

       Fig. 10

       Nucleotide DNA sequence (with the respective amino acid sequence) inserted at the XbaI site of the pMRS30 expression

20   vector to give the pMRS172 expression vector. This DNA includes the sequence termed UBILacI (see fig. 6) fused to the sequence partially corresponding to nucleotides 1222-1497 of the EMBL sequence J05581 followed by two translation stop codons, TGA and TAA. The encoded polypeptide thus includes the amino acid

25   sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 1222-1497 of the EMBL sequence J05581.

       Fig. 11

       Nucleotide DNA sequence (with the respective amino acid

30   sequence) inserted at the XbaI site of the pMRS30 expression vector to give the pMRS176 expression vector. This DNA includes the sequence named UBILacI (see fig. 6) fused to the sequence partially corresponding to nucleotides 136-1497 of the EMBL sequence J05581 followed by two translation stop codons, TGA and

TAA. The encoded polypeptide thus includes the amino acid sequence reported in Fig. 6, fused to the sequence including the amino acids encoded by the fragment 136-1497 of the EMBL sequence J05581.

5   **Fig. 12**

Electrophoretic analysis on 1% agarose gel in 1X TBE. mRNA extracted from CHO, CD34+ dendritic cells and dendritic cells from PBMC, respectively, transfected with pMRS169, and subjected to RT-PCR reaction either with (lanes 4, 8, 12) or without 10 (lanes 5, 9, 13) Reverse Transcriptase. Molecular weight DNA marker (lane 1); internal negative controls (lanes 2, 6); internal positive controls (lanes 3, 7, 10, 11); positive control from Promega kit (lane 14).

      Fig. 13

15   Nucleotide sequence of the pMRS30 expression vector. The 1-2862 region corresponds to the AccI (location 504) - BamHI (location 3369) region of the pSV2CAT vector (EMBL M77788); the 2863-3721 region includes the human cytomegalovirus promoter (human cytomegalovirus major immediate-early gene enhancer); the 20 3722-4905 region includes several cloning sites, including XbaI (location 3727), and the processing signal of the rabbit beta-globin gene.

      <u>Detailed description of the invention</u>

A DNA plasmid pool encoding, in eukaryotic cells, fragments 25 of the MUC-1 human protein antigen was prepared. Constructs are based on the mammalian expression vector termed pMRS30, described in figure 13 and previously claimed in the Patent Application WO95/11982, and contain partial sequences of the MUC-1 cDNAs reported in the EMBL database with accession number 30 J05581. MUC-1 encoding DNA was fragmented so that each fragment represents a discrete portion, partially overlapping to the adjacent ones. Administration of a mix of such plasmids can cause different plasmids to transfect different APC cells at the administration site. Therefore such cells produce and process

discrete portions of the MUC-1 protein giving the related peptides. In those conditions, the occurring subdominant and cryptic peptides can also be presented in association with class I MHC molecules thus generating a cytotoxic immune response.

The present invention thus relates to the use of a group of four constructs (Figures 1 to 4) containing MUC-1 cDNA partial fragments in admixture containing at least two of them and a group of four constructs (Figures 7 to 10) containing MUC-1 cDNA partial fragment preceded by the DNA encoding a protein sequence containing Ubiquitin and an Escherichia coli Lac I portion (Figure 6) used separately or in admixture containing at least two of them.

The present invention relates also to the use of the construct (Figure 5) containing the almost complete sequence of the MUC-1 cDNA and the construct (Figure 11) containing the almost complete sequence of the MUC-1 cDNA preceded by the DNA encoding a protein sequence containing Ubiquitin and an Escherichia coli Lac I portion.

The mixture of the four constructs containing the partial fragments of the MUC-1 cDNA and the mixture of the four constructs containing the partial fragments of the MUC-1 cDNA preceded by the DNA encoding a protein sequence, containing Ubiquitin and an Escherichia coli Lac I portion, represents a preferred embodiment of the present invention.

Constructs according to the present invention can be used in the anti-tumor therapy of patient affected with tumors characterized by high MUC-1 expression.

Constructs described in the present invention were obtained as follows.

In the case of the first series of constructs, the fragments of the MUC-1 DNA were obtained by RT-PCR from BT20 cell line or by DNA partial chemical synthesis. Such fragments were then cloned into the pMRS30 expression vector and verified by sequencing.

In the case of the second series of constructs, the
fragments were obtained from the first series of constructs by a
PCR re-amplification. These fragments were then fused to the DNA
encoding the Ubiquitin (obtained by RT-PCR from MCF7 cell line
5   mRNA) and a partial lacI sequence (obtained by PCR from the
commercial vector pGEX). DNA sequences thus obtained were then
cloned in the pMRS30 expression vector and verified by
sequencing. For the intended therapeutic or prophylactic uses,
fragments or constructs according to the invention are suitably
10   formulated, using carriers and methods previously employed in
naked DNA vaccines, as described for example in The
Immunologist, 1994, 2:1; WO 90/11092, Proc. Natl. Acad. Sci.
U.S.A., 1986, 83, 9551; US 5580859; Immunology today 19 (1998),
89-97); Proc. Natl. Acad. Sci. U.S.A. 90 (1993), 11478-11482;
15   Nat. Med. 3 (1997), 526-532; Vaccine 12 (1994), 1495-1498; DNA
Cell. Biol. 12 (1993), 777-783. The dosages will be determined
on the basis of clinical and pharmacological-toxicological
trials. Generally speaking, they will be comprised between 0.005
µg/kg and 5 µg/kg of the fragment mix. The composition of the
20   invention can also contain a cytokine or a cytokine encoding
plasmid.

The invention will be further illustrated by means of the
following examples.

Example 1. Plasmid pMRS166 construction.

25   BT20 tumor cells (ATCC HTB-19) were cultured in Eagles MEM
supplemented with 10% fetal calf serum. Ten million cells were
trypsinized, washed with PBS, and mRNA extracted.

An aliquot of this RNA was subjected to RT-PCR (reverse
transcriptase-polymerase chain reaction) reaction in the
30   presence of the following synthetic oligonucleotides:

V11 (5 GATCTCTAGAATGACAGGTTCTGGTCATGCAAGC 3)

V4 (5 GATCTCTAGAAAGCTTATCAACCTGAAGCTGGTTCCGTGGC 3)

The produced DNA fragment, purified and digested with the
restriction enzyme XbaI, was cloned into the pMRS30 expression

vector, containing the human cytomegalovirus promoter and the beta-globin polyadenylation signal as claimed in the Patent WO9511982. The resulting pMRS166 vector contains a DNA fragment including the ATG codon, the sequence corresponding to the

5   nucleotides 136-339 of the EMBL sequence J05581, and two stop codons, TGA and TAA.

This fragment is reported in fig. 1.

Example 2. Plasmid pMRS169 construction.

An aliquot of the RNA obtained as reported in example 1 was

10   amplified by RT-PCR in the presence of the following synthetic oligonuclotides:

V12 (5 GATCTCTAGAATGGTGCCCAGCTCTACTGAGAAGAATGC 3)

V15 (5 GGCGGTGGAGCCCGGGGCTGGCTTGT 3)

The produced DNA fragment, purified and digested with the

15   restriction enzymes SmaI and XbaI, was fused, by the SmaI restriction site, to a DNA fragment entirely synthetically constructed, and including a sequence partially corresponding to the nucleotides 457-720 of the EMBL sequence J05581 and two stop codons, TGA and TAA. The whole fragment was thus cloned in the

20   XbaI site of the pMRS30 expression vector. The resulting pMRS169 vector contains a DNA fragment including the ATG codon, the sequence partially corresponding to the nucleotides 205-720 of the EMBL sequence J05581, and two stop codons, TGA and TAA.

This fragment is reported in fig. 2.

25   **Example 3. Plasmid pMRS168 construction.**

An aliquot of the RNA obtained as reported in example 1 was amplified by RT-PCR in the presence of the following synthetic oligonuclotides:

V13 (5 GATCTCTAGAATGGGCTCAGCTTCTACTCTGGTGCACAACGGC 3)

30   V8   (5 GATCTCTAGAAAGCTTATCACAAGGCAATGAGATAGACAATGGCC 3)

The produced DNA fragment, purified and digested with the restriction enzyme XbaI was cloned in the pMRS30 expression vector. The resulting pMRS168 vector contains a DNA fragment including the ATG codon, the sequence corresponding to the

13

nucleotides 631-1275 of the EMBL sequence J05581, and two stop codons, TGA and TAA.

This fragment is reported in fig. 3.

Example 4. Plasmid pMRS167 construction.

5 An aliquot of the RNA obtained as reported in example 1 was subjected to RT-PCR reaction in the presence of the following synthetic oligonucleotides:

V14 (5 GATCTCTAGAATGCTGGTGCTGGTCTGTGTTCTGGTTGCGC 3)

V10 (5 GATCTCTAGAAAGCTTATCACAAGTTGGCAGAAGTGGCTGC 3)

10 The produced DNA fragment, purified and digested with the restriction enzyme XbaI was cloned in the pMRS30 expression vector. The resulting pMRS167 vector contains a DNA fragment including the ATG codon, the sequence corresponding to the nucleotides 1222-1497 of the EMBL sequence J05581, and two stop 15 codons, TGA and TAA.

This fragment is reported in fig. 4.

Example 5. Plasmid pMRS175 construction.

pMRS166, 169, 168, 167 plasmids were subjected to PCR reaction in the presence of the following nucleotide pairs:

20 V11 (see example 1)

V18 (5 AACCTGAAGCTGGTTCCGTGGC 3) for pMRS166

V19 (5 GTGCCCAGCTCTACTGAGAAGAATGC 3)

V20 (5 GCTGGGAATTGAGAATGGAGTGCTCTTGC 3) for pMRS169

V21 (5 GGCTCAGCTTCTACTCTGGTGCACAACGGC 3)

25 V22 (5 CAAGGCAATGAGATAGACAATGGCC 3) for pMRS168

V23 (5 CTGGTGCTGGTCTGTGTTCTGGTTGCG 3)

V10 (see example 4) for pMRS167

The four DNA fragments obtained in the respective PCR reactions were mixed in equimolar amounts and PCR reacted in the 30 presence of the V11 and V10 oligonuclotides.

The produced DNA fragment, purified and digested with the XbaI restriction enzyme, was cloned in the pMRS30 expression vector. The resulting pMRS175 vector contains a DNA fragment including the ATG codon, the sequence partially corresponding to

14

the nucleotides 136-1497 of the EMBL sequence J05581 and two
stop codons TGA and TAA.

This fragment is reported in fig. 5.

Example 6. Plasmid pMRS171 construction.

5   MCF7 tumor cells (ATCC HTB-22) were cultured in Eagles MEM
supplemented with 10% fetal calf serum. Ten million cells were
trypsinized, washed with PBS, and mRNA extracted.

An aliquot of this RNA was subjected to RT-PCR in the
presence of the following synthetic oligonucleotides:

10   UBIup (5GATCTCTAGAATGCAGATCTTCGTGAAGACCCTGACTGGT 3)

UBIdown

(5TCACCAGCGAGACGGGCAACAGCCATGCACCACTACCGTGCCTCCCACCTCTGAGACGGAGC
ACCAGG 3)

The reaction produces a DNA fragment termed fragment 1.

15   DNA from pGEX11T (Pharmacia) was subjected to PCR reaction
in the presence of the following synthetic oligonucleotides:

LacIup (5CCTCCGTCTCAGAGGTGGGAGGCACGGTAGTGGTGCATGGCTGTTGCCC
GTCTCGCTGGTGAAAAG 3)

LacIdown (5GATCGGATCCTCGGGAAACCTGTCGTGCCAGCTGC 3)

20   This reaction gives a DNA fragment termed fragment 2.

The 1 and 2 DNA fragments, obtained in the respective PCR
reactions, were mixed in equimolar amounts and subjected to PCR
reaction in presence of the UBIup and LacIdown oligonucleotides.

The produced DNA fragment, purified and digested with the

25   restriction enzymes XbaI and BamHI, was cloned into the pUC18
commercial plasmid. The resulting pMRS156 vector contains a DNA
fragment including the sequence encoding the ubiquitin fused to
the sequence encoding a bacterial beta-galactosidase portion.
This fragment, termed UBILacI, is reported in fig. 6.

30   Plasmid pMRS166 DNA was subjected to a PCR reaction in
presence of the following synthetic oligonucleotides:

V3 (5GATCGGATCCACAGGTTCTGGTCATGCAAGC 3)

V4 (see Example 1)

The produced DNA fragment, purified and digested with the

restriction enzymes XbaI and BamHI, was fused, by ligation into
the two BamHI sites, to the UBILacI fragment deriving from the
pMRS156 plasmid. The resulting fragment was cloned into the
pMRS30 expression vector. The resulting pMRS171 vector contains

5   a DNA fragment including the UBILacI sequence, the sequence
corresponding to the 136-339 nucleotides of the EMBL sequence
J05581 and two stop codons, TGA and TAA. This fragment is
reported in fig. 7.

        Example 7. Plasmid pMRS174 construction.

10   Plasmid pMRS169 DNA was subjected to PCR reaction in the
presence of the following synthetic oligonucleotides:

        V5 (5GATCGGATCCGTGCCCAGCTCTACTGAGAAGAATGC 3)

        V6(5GATCTCTAGAAAGCTTATCAGCTGGGAATTGAGAATGGAGTGCTCTTGC 3)

        The produced DNA fragment, purified and digested with the

15   restriction enzymes XbaI and BamHI, was fused, by ligation into
the two BamHI sites, to the UBILacI fragment deriving from the
pMRS156 plasmid. The resulting fragment was cloned into the
pMRS30 expression vector. The resulting pMRS174 vector contains
a DNA fragment including the UBILacI sequence, the sequence

20   corresponding to the 205-720 nucleotides of the EMBL sequence
J05581, and two stop codons, TGA and TAA. This fragment is
reported in fig. 8.

        Example 8. Plasmid pMRS173 construction.

        Plasmid pMRS168 DNA was subjected to PCR reaction in the

25   presence of the following synthetic oligonucleotides:

        V7 (5GATCGGATCCGGCTCAGCTTCTACTCTGGTGCACAACGGC 3)

        V8 (see example 3)

        The produced DNA fragment, purified and digested with the
restriction enzymes XbaI and BamHI, was fused, by ligation into

30   the two BamHI sites, to the UBILacI fragment deriving from the
pMRS156 plasmid. The resulting fragment was cloned into the
pMRS30 expression vector. The resulting pMRS173 vector contains
a DNA fragment including the UBILacI sequence, the sequence
corresponding to the 631-1275 nucleotides of the EMBL sequence

J05581, and two stop codons, TGA and TAA. This fragment is reported in fig. 9.

Example 9. Plasmid pMRS172 construction.

Plasmid pMRS167 DNA was subjected to PCR reaction in the presence of the following synthetic oligonucleotides:

V9 (5 GATCGGATCCCTGGTGCTGGTCTGTGTTCTGGTTGCGC 3)

V10 (see example 4)

The produced DNA fragment, purified and digested with the restriction enzymes XbaI and BamHI, was fused, by ligation into the two BamHI sites, to the UBILacI fragment deriving from pMRS156 plasmid. The resulting fragment was cloned into the pMRS30 expression vector. The resulting pMRS172 vector contains a DNA fragment including the UBILacI sequence, the sequence corresponding to the 1222-1497 nucleotides of the EMBL sequence J05581, and two stop codons, TGA and TAA. This fragment is reported in fig. 10.

Example 10. Plasmid pMRS176 construction.

Plasmid pMRS167 DNA was subjected PCR reaction in the presence of the following synthetic oligonucleotides:

V3 (see example 6)

V10 (see example 4)

The produced DNA fragment, purified and digested with the restriction enzymes XbaI and BamHI, was fused, by ligation into the two BamHI sites, to the UBILacI fragment deriving from pMRS156 plasmid. The resulting fragment was cloned into the pMRS30 expression vector. The resulting pMRS176 vector contains a DNA fragment including the UBILacI sequence, the sequence corresponding to the 136-1497 nucleotides of the EMBL sequence J05581, and two stop codons, TGA and TAA. This fragment is reported in fig. 11.

Example 11. Eukaryotic cell transfection and testing for transcription.

CHO (Chinese Hamster Ovary) cells were cultured in alpha MEM supplemented with ribonucleotides and deoxyribonucleotides

at transfection time.

Dendritic cells were obtained from CD34+ hemopoietic precursors cultured in IMDM without serum, supplemented with GM-CSF, IL4, SCF, Flt3 and TNFalpha. After 7 days the obtained cell

5   population was transfected.

Dendritic cells were obtained from monocytes isolated from PBMC (peripheral blood mononuclear cells), cultured in RPMI supplemented with FCS, GM-CSF, and IL-4. After 7 days the obtained cell population was transfected.

10   In each case, about one million cells were transfected with one of the plasmids reported in examples 1 to 10. Transfection was carried out using 3 μg of plasmid DNA and 4 μl of DMRIE (Gibco) by lipofection.

After 24 hours cells were harvested, washed with PBS and

15   lysed in order to extract the mRNA.

A mRNA aliquot was subjected to RT-PCR reaction in the presence of the oligonucleotide pair specific for the transfected DNA plasmid.

This experiment was carried out for each plasmid reported

20   in the examples 1 to 10, using the following oligonucleotide pairs: V11/V4 for pMRS166, V12/V6 for pMRS169, V13/V8 for pMRS168, V4/V10 for pMRS167, V4/V10 for pMRS175, UBIup/V4 for pMRS171, UBIup/V6 for pMRS174, UBIup/V8 for pMRS173, UBIup/V10 for pMRS172, V14/V10 for pMRS176.

25   As a representative example, figure 12 reports the electrophoretic analysis of the DNA fragments obtained by RT-PCR from the mRNA of the three cell populations, transfected with the pMRS169 plasmid. In this case the oligonucleotide pair V12/V6 was used.

30   Example 12. *In vivo* study results.

In the *in vivo* studies, the mixtures of the four fragments and the pMRS30 plasmid (vector without insert and thus used as a negative control) were used. In order to test the occurred immunization, an ELISA test was used to show the human mucin

specific antigens.

The *in vivo* studies were conducted using human MUC1 transgenic C57BL mice. As a consequence in these animals the MUC1 protein represents a self-protein. The employed vaccination schedule consists of 3 intradermic (dorsal portion, 50 micrograms DNA for each side) administrations (at days 0, 14, 28) of 100 micrograms plasmid DNA. At day 14 after the last administration, the animals were sacrificed and sera were tested for anti-human mucin antibodies.

The assayed fragment mixes, object of the present invention, stimulated a good immune response in the treated animals.

On the other hand, vaccination experiments with a 60-aminoacid peptide corresponding to the 20 aminoacids reported in fig. 2, from location 86 to location 105, repeated three times (this peptide is termed 3XTR), were also carried out.

The two vaccinations differ in the type of the elicited antibody response. The antibody titer results much more higher in the vaccination with 3XTR. Furthermore the noticed IgG subtypes are in favor of an essentially humoral (antibody) response in the case of vaccination with 3XTR, and of a cellular response (cytotoxic) in the case of vaccination with DNA. For anti-tumor therapy, a principally cytotoxic immune response is preferable. Because the experiments were carried out on transgenic mice, in whom the human mucin is "self", we can foresee a similar response in humans. This response could justify the use, as DNA vaccines, of the compounds of the present invention in the treatment of MUC1 overexpressing human tumors.

CLAIMS

1. Pharmaceutical composition containing one or more DNA molecules, encoding fragments of a protein overexpressed in tumor cells in order to induce an antitumor Ag-specific immune response, in combination with suitable excipients and adjuvants.

2. Pharmaceutical composition according to claim 1 wherein the overexpressed protein is MUC-1.

3. Pharmaceutical composition according to claim 1 or 2 containing at least two DNA molecules each containing a cDNA sequence encoding a Mucin fragment (MUC-1).

4. Composition according to claim 3 containing at least three DNA molecules each containing a cDNA sequence encoding a Mucin fragment (MUC-1).

5. Composition according to claim 4 containing at least four DNA molecules each containing a cDNA sequence encoding a Mucin fragment (MUC-1).

6. Composition according to claims 3, 4 or 5 wherein the DNA sequences comprise about 200 to about 700 nucleotides, each sequence being contiguous and possibly partially overlapping, from about 50 to about 150 nucleotides at the 3' and/or 5' end, to the adjacent one.

7. Pharmaceutical composition according to any claim from 2 to 6 wherein the used mixture consists of, at least, two plasmid DNA molecules, each containing a DNA fragment selected from those whose sequences are described in figures 1, 2, 3, and 4.

8. Pharmaceutical composition according to claim 7 wherein the used mixture consists of the pool of plasmid DNA molecules, where each molecule contains a DNA fragment selected from those whose sequences are described in figures 1, 2, 3, and 4.

9. Pharmaceutical composition according to claim 1 or 2 wherein a plasmid DNA molecule containing the sequence described in figure 5 is used.

10. Pharmaceutical composition according to claims 7, 8, or 9

wherein the used plasmid DNA molecules derive from the fusion of
the pMRS30 expression vector in Fig. 13 to each sequence
described in figures 1, 2, 3, 4, 5.

11. Pharmaceutical composition according to claims 2 to 6
wherein the used sequences, corresponding to single fragments of
the protein, are preceded in the 5' termini by the sequence
described in Fig. 6 encoding the ubiquitin and a LacI portion
from Escherichia Coli.

12. Pharmaceutical composition according to claim 11 wherein the
mixture consists of one or more sequences deriving from joining
the pMRS30 expression vector, described in Fig. 13, to a DNA
sequence selected from those described in figures 7, 8, 9, and
10.

13. Pharmaceutical composition according to claim 11 wherein the
mixture consists of the totality of the sequences deriving from
joining the pMRS30 expression vector to a DNA sequence selected
from those described in figures 7, 8, 9, and 10.

14. Pharmaceutical composition according to claim 11 wherein the
mixture consists of a sequence deriving from joining the pMRS30
expression vector to the sequence described in figure 11.

15. Pharmaceutical composition according to any preceding
claims, further containing a cytokine or a cytokine encoding
plasmid.

16. A plasmid DNA molecule consisting of the pMRS30 expression
vector joined to a DNA sequence, encoding a MUC-1 protein
fragment and whose sequence is selected from the group of those
described in figures 1, 2, 3, 4, and 5.

17. A DNA molecule encoding a protein MUC-1 fragment preceded in
its 5' terminus by the sequence described in Fig. 6.

18. A DNA molecule according to claim 17 selected from those
described in figures 7, 8, 9, 10, and 11.

19. A plasmid DNA molecule obtained by joining the pMRS
expression vector to a DNA molecule selected from those of claim
17 or 18.

20. Use of DNA molecules of claims 16-19 in the preparation of a composition with anti-tumor effect.

1/19

Figure 1

```
  1  ATGACAGGTTCTGGTCATGCAAGCTCTACCCCAGGTGGAGAAAAG
  1▶ Met Thr Gly Ser Gly His Ala Ser Ser Thr Pro Gly Gly Glu Lys
 46  GAGACTTCGGCTACCCAGAGAAGTTCAGTGCCCAGCTCTACTGAG
 16▶ Glu Thr Ser Ala Thr Gln Arg Ser Ser Val Pro Ser Ser Thr Glu
 91  AAGAATGCTGTGAGTATGACCAGCAGCGTACTCTCCAGCCACAGC
 31▶ Lys Asn Ala Val Ser Met Thr Ser Ser Val Leu Ser Ser His Ser
136  CCCGGTTCAGGCTCCTCCACCACTCAGGGACAGGATGTCACTCTG
 46▶ Pro Gly Ser Gly Ser Ser Thr Thr Gln Gly Gln Asp Val Thr Leu
181  GCCCCGGCCACGGAACCAGCTTCAGGTTGATAA
 61▶ Ala Pro Ala Thr Glu Pro Ala Ser Gly • • • • • •
```

2/19

Figure 2

```
  1  ATGGTGCCCAGCTCTACTGAGAAGAATGCTGTGAGTATGACCAGC
  1▶MetValProSerSerThrGluLysAsnAlaValSerMetThrSer
 46  AGCGTACTCTCCAGCCACAGCCCCGGTTCAGGCTCCTCCACCACT
 16▶SerValLeuSerSerHisSerProGlySerGlySerSerThrThr
 91  CAGGGACAGGATGTCACTCTGGCCCCGGCCACGGAACCAGCTTCA
 31▶GlnGlyGlnAspValThrLeuAlaProAlaThrGluProAlaSer
136  GGTTCAGCTGCCACCTGGGGACAGGATGTCACCTCGGTCCCAGTC
 46▶GlySerAlaAlaThrTrpGlyGlnAspValThrSerValProVal
181  ACCAGGCCAGCCCTGGGCTCCACCACCCCGCCAGCCCACGATGTC
 61▶ThrArgProAlaLeuGlySerThrThrProProAlaHisAspVal
226  ACCTCAGCCCCGGACAACAAGCCAGCCCCGGGAAGTACTGCTCCA
 76▶ThrSerAlaProAspAsnLysProAlaProGlySerThrAlaPro
271  CCAGCACACGGTGTTACCTCGGCTCCGGATACCAGGCCGGCCCCA
 91▶ProAlaHisGlyValThrSerAlaProAspThrArgProAlaPro
316  GGTAGTACCGCCCCTCCTGCCCATGGTGTCACATCTGCCCCGGAC
106▶GlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp
361  AACAGGCCTGCATTGGGTAGTACAGCACCGCCAGTACACAACGTT
121▶AsnArgProAlaLeuGlySerThrAlaProProValHisAsnVal
406  ACTAGTGCCTCAGGCTCTGCTAGCGGCTCAGCTTCTACTCTGGTG
136▶ThrSerAlaSerGlySerAlaSerGlySerAlaSerThrLeuVal
451  CACAACGGCACCTCTGCGCGCGCGACCACAACCCCAGCGAGCAAG
151▶HisAsnGlyThrSerAlaArgAlaThrThrThrProAlaSerLys
496  AGCACTCCATTCTCAATTCCCAGCTGATAA
166▶SerThrProPheSerIleProSer••••••
```

Figure 3

```
  1  ATGGGCTCAGCTTCTACTCTGGTGCACAACGGCACCTCTGCCAGG
  1▶ MetGlySerAlaSerThrLeuValHisAsnGlyThrSerAlaArg
 46  GCTACCACAACCCCAGCCAGCAAGAGCACTCCATTCTCAATTCCC
 16▶ AlaThrThrThrProAlaSerLysSerThrProPheSerIlePro
 91  AGCCACCACTCTGATACTCCTACCACCCTTGCCAGCCATAGCACC
 31▶ SerHisHisSerAspThrProThrThrLeuAlaSerHisSerThr
136  AAGACTGATGCCAGTAGCACTCACCATAGCACGGTACCTCCTCTC
 46▶ LysThrAspAlaSerSerThrHisHisSerThrValProProLeu
181  ACCTCCTCCAATCACAGCACTTCTCCCCAGTTGTCTACTGGGGTC
 61▶ ThrSerSerAsnHisSerThrSerProGlnLeuSerThrGlyVal
226  TCTTTCTTTTTCCTGTCTTTTCACATTTCAAACCTCCAGTTTAAT
 76▶ SerPhePhePheLeuSerPheHisIleSerAsnLeuGlnPheAsn
271  TCCTCTCTGGAAGATCCCAGCACCGACTACTACCAAGAGCTGCAG
 91▶ SerSerLeuGluAspProSerThrAspTyrTyrGlnGluLeuGln
316  AGAGACATTTCTGAAATGTTTTTGCAGATTTATAAACAAGGGGGT
106▶ ArgAspIleSerGluMetPheLeuGlnIleTyrLysGlnGlyGly
361  TTTCTGGGCCTCTCCAATATTAAGTTCAGGCCAGGATCTGTGGTG
121▶ PheLeuGlyLeuSerAsnIleLysPheArgProGlySerValVal
406  GTACAATTGACTCTGGCCTTCCGAGAAGGTACCATCAATGTCCAC
136▶ ValGlnLeuThrLeuAlaPheArgGluGlyThrIleAsnValHis
451  GACGTGGAGACACAGTTCAATCAGTATAAAACGGAAGCAGCCTCT
151▶ AspValGluThrGlnPheAsnGlnTyrLysThrGluAlaAlaSer
496  CGATATAACCTGACGATCTCAGACGTCAGCGTGAGTGATGTGCCA
166▶ ArgTyrAsnLeuThrIleSerAspValSerValSerAspValPro
541  TTTCCTTTCTCTGCCCAGTCTGGGGCTGGGGTGCCAGGCTGGGGC
181▶ PheProPheSerAlaGlnSerGlyAlaGlyValProGlyTrpGly
586  ATCGCGCTGCTGGTGCTGGTCTGTGTTCTGGTTGCGCTGGCCATT
196▶ IleAlaLeuLeuValLeuValCysValLeuValAlaLeuAlaIle
631  GTCTATCTCATTGCCTTGTGATAA
211▶ ValTyrLeuIleAlaLeu••••••
```

4/19

Figure 4

```
  1  ATGCTGGTGCTGGTCTGTGTTCTGGTTGCGCTGGCCATTGTCTAT
  1▶ Met LeuVal LeuVal CysVal LeuVal Al aLeuAl a l l eVal Tyr
 46  CTCATTGCCTTGGCTGTCTGTCAGTGCCGCCGAAAGAACTACGGG
 16▶ Leu l l eAl aLeuAl aVal CysGl nCysArgArgLysAsnTyrGl y
 91  CAGCTGGACATCTTTCCAGCCCGGGATACCTACCATCCTATGAGC
 31▶ Gl nLeuAsp l l ePheProAl aArgAspThr TyrHi s ProMet Ser
136  GAGTACCCCACCTACCACACCCATGGGCGCTATGTGCCCCCTAGC
 46▶ Gl uTyrProThr TyrHi sThr Hi sGl yA rgTyrVal ProProSer
181  AGTACCGATCGTAGCCCCTATGAGAAGGTTTCTGCAGGTAATGGT
 61▶ Ser ThrAspArgSer ProTyrGl uLysVal Ser Al aGl yAsnGl y
226  GGCAGCAGCCTCTCTTACACAAACCCAGCAGTGGCAGCCACTTCT
 76▶ Gl ySer Ser LeuSer TyrThrAsnProAl aVal Al aAl aThr Ser
271  GCCAACTTGTGATAA
 91▶ Al aAsnLeu• • • • • •
```

Figure 5

```
  1  ATGACAGGTTCTGGTCATGCAAGCTCTACCCCAGGTGGAGAAAAG
  1▶ Met Thr Gl ySer Gl yHi sAl aSer Ser Thr ProGl yGl yGl uLy s
 46  GAGACTTCGGCTACCCAGAGAAGTTCAGTGCCCAGCTCTACTGAG
 16▶ Gl uThr Ser Al aThr Gl nArgSer Ser Val ProSer Ser Thr Gl u
 91  AAGAATGCTGTGAGTATGACCAGCAGCGTACTCTCCAGCCACAGC
 31▶ LysAsnAl aVal Ser Met Thr Ser Ser Val LeuSer Ser Hi s Ser
136  CCCGGTTCAGGCTCCTCCACCACTCAGGGACAGGATGTCACTCTG
 46▶ ProGl ySer Gl ySer Ser Thr Thr Gl nGl yGl nAspVal Thr Leu
181  GCCCCGGCCACGGAACCAGCTTCAGGTTCAGCTGCCACCTGGGGA
 61▶ Al aProAl aThr Gl uProAl aSer Gl ySer Al aAl aThr TrpGl y
226  CAGGATGTCACCTCGGTCCCAGTCACCAGGCCAGCCCTGGGCTCC
 76▶ Gl nAspVal Thr Ser Val ProVal Thr ArgProAl aLeuGl ySer
271  ACCACCCCGCCAGCCCACGATGTCACCTCAGCCCCGGACAACAAG
 91▶ Thr Thr ProProAl aHi sAspVal Thr Ser Al aProAspAsnLy s
316  CCAGCCCCGGGAAGTACCGCTCCACCAGCACACGGTGTTACCTCG
106▶ ProAl aProGl ySer Thr Al aProProAl aHi sGl yVal Thr Ser
361  GCTCCGGATACCAGGCCGGCCCCAGGTAGTACCGCCCCTCCTGCC
121▶ Al aProAspThr ArgProAl aProGl ySer Thr Al aProProAl a
406  CATGGTGTCACATCTGCCCCGGACAACAGGCCTGCATTGGGTAGT
136▶ Hi sGl yVal Thr Ser Al aProAspAsnArgProAl aLeuGl ySer
451  ACAGCACCGCCAGTACACAACGTTACTAGTGCCTCAGGCTCTGCT
151▶ Thr Al aProProVal Hi sAsnVal Thr Ser Al aSer Gl ySer Al a
496  AGCGGCTCAGCTTCTACTCTGGTGCACAACGGCACCTCTGCGCGC
166▶ Ser Gl ySer Al aSer Thr LeuVal Hi sAsnGl yThr Ser Al aArg
541  GCGACCACAACCCCAGCGAGCAAGAGCACTCCATTCTCAATTCCC
181▶ Al aThr Thr Thr ProAl aSer LysSer Thr ProPheSer I l ePro
586  AGCCACCACTCTGATACTCCTACCACCCTTGCCAGCCATAGCACC
196▶ Ser Hi s Hi s Ser AspThr ProThr Thr LeuAl aSer Hi sSer Thr
631  AAGACTGATGCCAGTAGCACTCACCATAGCACGGTACCTCCTCTC
211▶ LysThr AspAl aSer Ser Thr Hi s Hi s Ser Thr Val ProProLeu
676  ACCTCCTCCAATCACAGCACTTCTCCCCAGTTGTCTACTGGGGTC
226▶ Thr Ser Ser AsnHi s Ser Thr Ser ProGl nLeuSer Thr Gl yVal
721  TCTTTCTTTTTCCTGTCTTTTCACATTTCAAACCTCCAGTTTAAT
241▶ Ser PhePhePheLeuSer PheHi s I l eSer AsnLeuGl nPheAsn
766  TCCTCTCTGGAAGATCCCAGCACCGACTACTACCAAGAGCTGCAG
256▶ Ser Ser LeuGl uAspProSer Thr AspTyrTyrGl nGl uLeuGl n
811  AGAGACATTTCTGAAATGTTTTTGCAGATTTATAAACAAGGGGGT
271▶ ArgAsp I l eSer Gl uMet PheLeuGl n I l eTyrLysGl nGl yGl y
856  TTTCTGGGCCTCTCCAATATTAAGTTCAGGCCAGGATCTGTGGTG
286▶ PheLeuGl yLeuSer Asn I l eLysPheArgProGl ySer Val Val
```

(Continued)

Figure 5 (continued)

```
 901  GTACAATTGACTCTGGCCTTCCGAGAAGGTACCATCAATGTCCAC
 301▶ ValGlnLeuThrLeuAlaPheArgGluGlyThrIleAsnValHis
 946  GACGTGGAGACACAGTTCAATCAGTATAAAACGGAAGCAGCCTCT
 316▶ AspValGluThrGlnPheAsnGlnTyrLysThrGluAlaAlaSer
 991  CGATATAACCTGACGATCTCAGACGTCAGCGTGAGTGATGTGCCA
 331▶ ArgTyrAsnLeuThrIleSerAspValSerValSerAspValPro
1036  TTTCCTTTCTCTGCCCAGTCTGGGGCTGGGGTGCCAGGCTGGGGC
 346▶ PheProPheSerAlaGlnSerGlyAlaGlyValProGlyTrpGly
1081  ATCGCGCTGCTGGTGCTGGTCTGTGTTCTGGTTGCGCTGGCCATT
 361▶ IleAlaLeuLeuValLeuValCysValLeuValAlaLeuAlaIle
1126  GTCTATCTCATTGCCTTGGCTGTCTGTCAGTGCCGCCGAAAGAAC
 376▶ ValTyrLeuIleAlaLeuAlaValCysGlnCysArgArgLysAsn
1171  TACGGGCAGCTGGACATCTTTCCAGCCCGGGATACCTACCATCCT
 391▶ TyrGlyGlnLeuAspIlePheProAlaArgAspThrTyrHisPro
1216  ATGAGCGAGTACCCCACCTACCACACCCATGGGCGCTATGTGCCC
 406▶ MetSerGluTyrProThrTyrHisThrHisGlyArgTyrValPro
1261  CCTAGCAGTACCGATCGTAGCCCCTATGAGAAGGTTTCTGCAGGT
 421▶ ProSerSerThrAspArgSerProTyrGluLysValSerAlaGly
1306  AATGGTGGCAGCAGCCTCTCTTACACAAACCCAGCAGTGGCAGCC
 436▶ AsnGlyGlySerSerLeuSerTyrThrAsnProAlaValAlaAla
1351  ACTTCTGCCAACTTGTGATAA
 451▶ ThrSerAlaAsnLeu••••••
```

Figure 6

```
  1  ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC
  1▶ MetGlnIlePheValLysThrLeuThrGlyLysThrIleThrLeu
 46  GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC
 16▶ GluValGluProSerAspThrIleGluAsnValLysAlaLysIle
 91  CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31▶ GlnAspLysGluGlyIleProProAspGlnGlnArgLeuIlePhe
136  GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46▶ AlaGlyLysGlnLeuGluAspGlyArgThrLeuSerAspTyrAsn
181  ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT
 61▶ IleGlnLysGluSerThrLeuHisLeuValLeuArgLeuArgGly
226  GGGAGGCACGGTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76▶ GlyArgHisGlySerGlyAlaTrpLeuLeuProValSerLeuVal
271  AAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCC
 91▶ LysArgLysThrThrLeuAlaProAsnThrGlnThrAlaSerPro
316  CGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
106▶ ArgAlaLeuAlaAspSerLeuMetGlnLeuAlaArgGlnValSer
361  CGAGGATCC
121▶ ArgGlySer
```

Figure 7

```
  1  ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC
  1▶ MetGlnIlePheValLysThrLeuThrGlyLysThrIleThrLeu
 46  GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC
 16▶ GluValGluProSerAspThrIleGluAsnValLysAlaLysIle
 91  CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31▶ GlnAspLysGluGlyIleProProAspGlnGlnArgLeuIlePhe
136  GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46▶ AlaGlyLysGlnLeuGluAspGlyArgThrLeuSerAspTyrAsn
181  ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT
 61▶ IleGlnLysGluSerThrLeuHisLeuValLeuArgLeuArgGly
226  GGGAGGCACGGTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76▶ GlyArgHisGlySerGlyAlaTrpLeuLeuProValSerLeuVal
271  AAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCC
 91▶ LysArgLysThrThrLeuAlaProAsnThrGlnThrAlaSerPro
316  CGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
106▶ ArgAlaLeuAlaAspSerLeuMetGlnLeuAlaArgGlnValSer
361  CGAGGATCCACAGGTTCTGGTCATGCAAGCTCTACCCCAGGTGGA
121▶ ArgGlySerThrGlySerGlyHisAlaSerSerThrProGlyGly
406  GAAAAGGAGACTTCGGCTACCCAGAGAAGTTCAGTGCCCAGCTCT
136▶ GluLysGluThrSerAlaThrGlnArgSerSerValProSerSer
451  ACTGAGAAGAATGCTGTGAGTATGACCAGCAGCGTACTCTCCAGC
151▶ ThrGluLysAsnAlaValSerMetThrSerSerValLeuSerSer
496  CACAGCCCCGGTTCAGGCTCCTCCACCACTCAGGGACAGGATGTC
166▶ HisSerProGlySerGlySerSerThrThrGlnGlyGlnAspVal
541  ACTCTGGCCCCGGCCACGGAACCAGCTTCAGGTTGATAA
181▶ ThrLeuAlaProAlaThrGluProAlaSerGly******
```

Figure 8

```
  1  ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC
  1▶ Met Gl n I l ePheVa l LysThr LeuThr Gl yLysThr I l eThr Leu
 46  GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC
 16▶ Gl uVa l Gl uP roSer A spThr I l eGl uAsnVa l LysA l aLys I l e
 91  CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31▶ Gl nAspLys Gl uGl y I l eP roP roAspGl nGl nArgLeu I l ePhe
136  GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46▶ A l aGl yLysGl nLeuGl uAspGl yA rgThr LeuSer A spTyrAsn
181  ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT
 61▶ I l eGl nLys Gl uSer Thr LeuHi s LeuVa l LeuArgLeuArgGl y
226  GGGAGGCACGGTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76▶ Gl yA rgHi s Gl ySer Gl yA l aT rpLeuLeuP roVa l Ser LeuVa l
271  AAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCC
 91▶ LysArgLysThr Thr LeuA l aP roAsnThr Gl nThr A l aSer Pro
316  CGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
106▶ A rgA l aLeuA l aAspSer LeuMet Gl nLeuA l aArgGl nVa l Ser
361  CGAGGATCCGTGCCCAGCTCTACTGAGAAGAATGCTGTGAGTATG
121▶ A rgGl ySer Va l P roSer Ser Thr Gl uLysAsnA l aVa l Ser Met
406  ACCAGCAGCGTACTCTCCAGCCACAGCCCCGGTTCAGGCTCCTCC
136▶ Thr Ser Ser Va l LeuSer Ser Hi s Ser P roGl ySer Gl ySer Ser
451  ACCACTCAGGGACAGGATGTCACTCTGGCCCCGGCCACGGAACCA
151▶ Thr Thr Gl nGl yGl nAspVa l Thr LeuA l aP roA l aThr Gl uPro
496  GCTTCAGGTTCAGCTGCCACCTGGGGACAGGATGTCACCTCGGTC
166▶ A l aSer Gl ySer A l aA l aThr T rpGl yGl nAspVa l Thr Ser Va l
541  CCAGTCACCAGGCCAGCCCTGGGCTCCACCACCCCGCCAGCCCAC
181▶ P roVa l Thr A rgP roA l aLeuGl ySer Thr Thr P roP roA l aHi s
586  GATGTCACCTCAGCCCCGGACAACAAGCCAGCCCCGGGAAGTACT
196▶ AspVa l Thr Ser A l aP roAspAsnLys P roA l aP roGl ySer Thr
631  GCTCCACCAGCACACGGTGTTACCTCGGCTCCGGATACCAGGCCG
211▶ A l aP roP roA l aHi s Gl yVa l Thr Ser A l aP roAspThr A rgPro
676  GCCCCAGGTAGTACCGCCCCTCCTGCCCATGGTGTCACATCTGCC
226▶ A l aP roGl ySer Thr A l aP roP roA l aHi s Gl yVa l Thr Ser A l a
721  CCGGACAACAGGCCTGCATTGGGTAGTACAGCACCGCCAGTACAC
241▶ P roAspAsnArgP roA l aLeuGl ySer Thr A l aP roP roVa l Hi s
766  AACGTTACTAGTGCCTCAGGCTCTGCTAGCGGCTCAGCTTCTACT
256▶ AsnVa l Thr Ser A l aSer Gl ySer A l aSer Gl ySer A l aSer Thr
811  CTGGTGCACAACGGCACCTCTGCGCGCGCGACCACAACCCCAGCG
271▶ LeuVa l Hi s AsnGl yThr Ser A l aA rgA l aThr Thr Thr P roA l a
856  AGCAAGAGCACTCCATTCTCAATTCCCAGCTGATAA
286▶ Ser LysSer Thr P roPheSer I l eP roSer • • • • • •
```

Figure 9

```
  1  ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC
  1▶ MetGl n I I ePheVa I LysThr LeuThr GI yLysThr I I eThr Leu
 46  GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC
 16▶ GI uVa I GI uProSer AspThr I I eGI uAsnVa I LysA I aLys I I e
 91  CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31▶ GI nAspLysGI uGI y I I eProP roAspGI nGI nArgLeu I I ePhe
136  GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46▶ A I aGI yLysGI nLeuGI uAspGI yA rgThr LeuSer AspTyrAsn
181  ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT
 61▶ I I eGI nLysGI uSer Thr LeuHi sLeuVa I LeuArgLeuArgGI y
226  GGGAGGCACGGTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76▶ GI yA rgHi sGI ySer GI yA I aT rpLeuLeuProVa I Ser LeuVa I
271  AAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCC
 91▶ LysArgLysThr Thr LeuA I aP roAsnThr GI nThr A I aSer Pro
316  CGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
106▶ A rgA I aLeuA I aAspSer LeuMe t GI nLeuA I aArgGI nVa I Ser
361  CGAGGATCCGGCTCAGCTTCTACTCTGGTGCACAACGGCACCTCT
121▶ A rgGI ySer GI ySer A I aSer Thr LeuVa I Hi sAsnGI yThr Ser
406  GCCAGGGCTACCACAACCCCAGCCAGCAAGAGCACTCCATTCTCA
136▶ A I aArgA I aThr Thr Thr P roA I aSer LysSer Thr P roPheSer
451  ATTCCCAGCCACCACTCTGATACTCCTACCACCCTTGCCAGCCAT
151▶ I I eP roSer Hi sHi sSer AspThr P roThr Thr LeuA I aSer Hi s
496  AGCACCAAGACTGATGCCAGTAGCACTCACCATAGCACGGTACCT
166▶ Ser Thr LysThrAspA I aSer Ser Thr Hi sHi sSer Thr Va I P ro
541  CCTCTCACCTCCTCCAATCACAGCACTTCTCCCCAGTTGTCTACT
181▶ ProLeuThr Ser Ser AsnHi sSer Thr Ser P roGI nLeuSer Thr
586  GGGGTCTCTTTCTTTTTCCTGTCTTTTCACATTTCAAACCTCCAG
196▶ GI yVa I Ser PhePhePheLeuSer PheHi s I I eSer AsnLeuGI n
631  TTTAATTCCTCTCTGGAAGATCCCAGCACCGACTACTACCAAGAG
211▶ PheAsnSer Ser LeuGI uAspP roSer ThrAspTyrTyrGI nGI u
676  CTGCAGAGAGACATTTCTGAAATGTTTTTGCAGATTTATAAACAA
226▶ LeuGI nArgAsp I I eSer GI uMet PheLeuGI n I I eTyrLysGI n
721  GGGGGTTTTCTGGGCCTCTCCAATATTAAGTTCAGGCCAGGATCT
241▶ GI yGI yPheLeuGI yLeuSer Asn I I eLysPheArgP roGI ySer
766  GTGGTGGTACAATTGACTCTGGCCTTCCGAGAAGGTACCATCAAT
256▶ Va I Va I Va I GI nLeuThr LeuA I aPheArgGI uGI yThr I I eAsn
811  GTCCACGACGTGGAGACACAGTTCAATCAGTATAAAACGGAAGCA
271▶ Va I Hi sAspVa I GI uThr GI nPheAsnGI nTyrLysThr GI uA I a
856  GCCTCTCGATATAACCTGACGATCTCAGACGTCAGCGTGAGTGAT
286▶ A I aSer A rgTyrAsnLeuThr I I eSer AspVa I Ser Va I Ser Asp
901  GTGCCATTTCCTTTCTCTGCCCAGTCTGGGGCTGGGGTGCCAGGC
301▶ Va I P roPheP roPheSer A I aGI nSer GI yA I aGI yVa I P roGI y
946  TGGGGCATCGCGCTGCTGGTGCTGGTCTGTGTTCTGGTTGCGCTG
316▶ T rpGI y I I eA I aLeuLeuVa I LeuVa I CysVa I LeuVa I A I aLeu
991  GCCATTGTCTATCTCATTGCCTTGTGATAA
331▶ A I a I I eVa I TyrLeu I I eA I aLeu • • • • •
```

Figure 10

```
  1  ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC
  1▶ MetGlnIlePheValLysThrLeuThrGlyLysThrIleThrLeu
 46  GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC
 16▶ GluValGluProSerAspThrIleGluAsnValLysAlaLysIle
 91  CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31▶ GlnAspLysGluGlyIleProProAspGlnGlnArgLeuIlePhe
136  GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46▶ AlaGlyLysGlnLeuGluAspGlyArgThrLeuSerAspTyrAsn
181  ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT
 61▶ IleGlnLysGluSerThrLeuHisLeuValLeuArgLeuArgGly
226  GGGAGGCACGGTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76▶ GlyArgHisGlySerGlyAlaTrpLeuLeuProValSerLeuVal
271  AAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCC
 91▶ LysArgLysThrThrLeuAlaProAsnThrGlnThrAlaSerPro
316  CGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
106▶ ArgAlaLeuAlaAspSerLeuMetGlnLeuAlaArgGlnValSer
361  CGAGGATCCCTGGTGCTGGTCTGTGTTCTGGTTGCGCTGGCCATT
121▶ ArgGlySerLeuValLeuValCysValLeuValAlaLeuAlaIle
406  GTCTATCTCATTGCCTTGGCTGTCTGTCAGTGCCGCCGAAAGAAC
136▶ ValTyrLeuIleAlaLeuAlaValCysGlnCysArgArgLysAsn
451  TACGGGCAGCTGGACATCTTTCCAGCCCGGGATACCTACCATCCT
151▶ TyrGlyGlnLeuAspIlePheProAlaArgAspThrTyrHisPro
496  ATGAGCGAGTACCCCACCTACCACACCCATGGGCGCTATGTGCCC
166▶ MetSerGluTyrProThrTyrHisThrHisGlyArgTyrValPro
541  CCTAGCAGTACCGATCGTAGCCCCTATGAGAAGGTTTCTGCAGGT
181▶ ProSerSerThrAspArgSerProTyrGluLysValSerAlaGly
586  AATGGTGGCAGCAGCCTCTCTTACACAAACCCAGCAGTGGCAGCC
196▶ AsnGlyGlySerSerLeuSerTyrThrAsnProAlaValAlaAla
631  ACTTCTGCCAACTTGTGATAA
211▶ ThrSerAlaAsnLeu••••••
```

12/19

Figure 11

```
  1  ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC
  1▶ MetGlnIlePheValLysThrLeuThrGlyLysThrIleThrLeu
 46  GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC
 16▶ GluValGluProSerAspThrIleGluAsnValLysAlaLysIle
 91  CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTCATCTTT
 31▶ GlnAspLysGluGlyIleProProAspGlnGlnArgLeuIlePhe
136  GCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAAC
 46▶ AlaGlyLysGlnLeuGluAspGlyArgThrLeuSerAspTyrAsn
181  ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT
 61▶ IleGlnLysGluSerThrLeuHisLeuValLeuArgLeuArgGly
226  GGGAGGCACGGTAGTGGTGCATGGCTGTTGCCCGTCTCGCTGGTG
 76▶ GlyArgHisGlySerGlyAlaTrpLeuLeuProValSerLeuVal
271  AAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCC
 91▶ LysArgLysThrThrLeuAlaProAsnThrGlnThrAlaSerPro
316  CGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
106▶ ArgAlaLeuAlaAspSerLeuMetGlnLeuAlaArgGlnValSer
361  CGAGGATCCACAGGTTCTGGTCATGCAAGCTCTACCCCAGGTGGA
121▶ ArgGlySerThrGlySerGlyHisAlaSerSerThrProGlyGly
406  GAAAAGGAGACTTCGGCTACCCAGAGAAGTTCAGTGCCCAGCTCT
136▶ GluLysGluThrSerAlaThrGlnArgSerSerValProSerSer
451  ACTGAGAAGAATGCTGTGAGTATGACCAGCAGCGTACTCTCCAGC
151▶ ThrGluLysAsnAlaValSerMetThrSerSerValLeuSerSer
496  CACAGCCCCGGTTCAGGCTCCTCCACCACTCAGGGACAGGATGTC
166▶ HisSerProGlySerGlySerSerThrThrGlnGlyGlnAspVal
541  ACTCTGGCCCCGGCCACGGAACCAGCTTCAGGTTCAGCTGCCACC
181▶ ThrLeuAlaProAlaThrGluProAlaSerGlySerAlaAlaThr
586  TGGGGACAGGATGTCACCTCGGTCCCAGTCACCAGGCCAGCCCTG
196▶ TrpGlyGlnAspValThrSerValProValThrArgProAlaLeu
631  GGCTCCACCACCCCGCCAGCCCACGATGTCACCTCAGCCCCGGAC
211▶ GlySerThrThrProProAlaHisAspValThrSerAlaProAsp
676  AACAAGCCAGCCCCGGGAAGTACCGCTCCACCAGCACACGGTGTT
226▶ AsnLysProAlaProGlySerThrAlaProProAlaHisGlyVal
721  ACCTCGGCTCCGGATACCAGGCCGGCCCCAGGTAGTACCGCCCCT
241▶ ThrSerAlaProAspThrArgProAlaProGlySerThrAlaPro
766  CCTGCCCATGGTGTCACATCTGCCCCGGACAACAGGCCTGCATTG
256▶ ProAlaHisGlyValThrSerAlaProAspAsnArgProAlaLeu
811  GGTAGTACAGCACCGCCAGTACACAACGTTACTAGTGCCTCAGGC
271▶ GlySerThrAlaProProValHisAsnValThrSerAlaSerGly
856  TCTGCTAGCGGCTCAGCTTCTACTCTGGTGCACAACGGCACCTCT
286▶ SerAlaSerGlySerAlaSerThrLeuValHisAsnGlyThrSer
```

(Continued)

13/19

Figure 11 (continued)

```
 901  GCGCGCGCGACCACAACCCCAGCGAGCAAGAGCACTCCATTCTCA
 301▶ AlaArgAlaThr Thr Thr ProAlaSer LysSer Thr ProPheSer
 946  ATTCCCAGCCACCACTCTGATACTCCTACCACCCTTGCCAGCCAT
 316▶ IleProSer HisHisSerAspThr ProThr Thr LeuAlaSer His
 991  AGCACCAAGACTGATGCCAGTAGCACTCACCATAGCACGGTACCT
 331▶ Ser Thr LysThrAspAlaSer Ser Thr HisHisSer Thr Val Pro
1036  CCTCTCACCTCCTCCAATCACAGCACTTCTCCCCAGTTGTCTACT
 346▶ ProLeuThr Ser SerAsnHisSer Thr Ser ProGlnLeuSer Thr
1081  GGGGTCTCTTTCTTTTTCCTGTCTTTTCACATTTCAAACCTCCAG
 361▶ GlyVal Ser PhePhePheLeuSer PheHis IleSerAsnLeuGln
1126  TTTAATTCCTCTCTGGAAGATCCCAGCACCGACTACTACCAAGAG
 376▶ PheAsnSer Ser LeuGluAspProSer ThrAspTyrTyrGlnGlu
1171  CTGCAGAGAGACATTTCTGAAATGTTTTTGCAGATTTATAAACAA
 391▶ LeuGlnArgAspIleSer GluMet PheLeuGlnIleTyrLysGln
1216  GGGGGTTTTCTGGGCCTCTCCAATATTAAGTTCAGGCCAGGATCT
 406▶ GlyGlyPheLeuGlyLeuSerAsnIleLysPheArgProGlySer
1261  GTGGTGGTACAATTGACTCTGGCCTTCCGAGAAGGTACCATCAAT
 421▶ Val Val Val GlnLeuThr LeuAlaPheArgGluGlyThr IleAsn
1306  GTCCACGACGTGGAGACACAGTTCAATCAGTATAAAACGGAAGCA
 436▶ Val HisAspVal GluThr GlnPheAsnGlnTyrLysThr GluAla
1351  GCCTCTCGATATAACCTGACGATCTCAGACGTCAGCGTGAGTGAT
 451▶ AlaSer ArgTyrAsnLeuThr IleSerAspVal Ser Val SerAsp
1396  GTGCCATTTCCTTTCTCTGCCCAGTCTGGGGCTGGGGTGCCAGGC
 466▶ Val ProPheProPheSer AlaGlnSer GlyAlaGlyVal ProGly
1441  TGGGGCATCGCGCTGCTGGTGCTGGTCTGTGTTCTGGTTGCGCTG
 481▶ TrpGly IleAlaLeuLeuVal LeuVal CysVal LeuVal AlaLeu
1486  GCCATTGTCTATCTCATTGCCTTGGCTGTCTGTCAGTGCCGCCGA
 496▶ Ala IleVal TyrLeu IleAlaLeuAlaVal CysGlnCysArgArg
1531  AAGAACTACGGGCAGCTGGACATCTTTCCAGCCCGGGATACCTAC
 511▶ LysAsnTyrGlyGlnLeuAspIlePheProAlaArgAspThr Tyr
1576  CATCCTATGAGCGAGTACCCCACCTACCACACCCATGGGCGCTAT
 526▶ HisProMet Ser GluTyrProThr TyrHisThr HisGlyArgTyr
1621  GTGCCCCCTAGCAGTACCGATCGTAGCCCCTATGAGAAGGTTTCT
 541▶ Val ProProSer Ser ThrAspArgSer ProTyrGluLysVal Ser
1666  GCAGGTAATGGTGGCAGCAGCCTCTCTTACACAAACCCAGCAGTG
 556▶ AlaGlyAsnGlyGlySer Ser LeuSer TyrThrAsnProAlaVal
1711  GCAGCCACTTCTGCCAACTTGTGATAA
 571▶ AlaAlaThr Ser AlaAsnLeu••••••
```
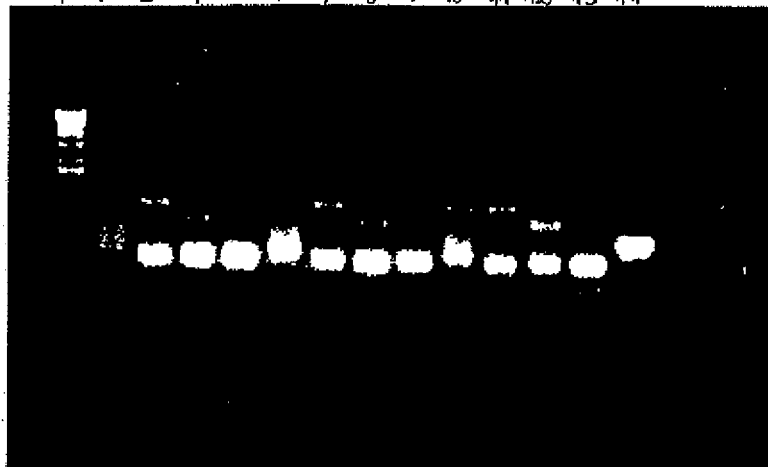
Figure 13

```
   1   CCAGGAAGCTCCTCTGTGTCCTCATAAACCCTAACCTCCTCTACTTGAGA

  51   GGACATTCCAATCATAGGCTGCCCATCCACCCTCTGTGTCCTCCTGTTAA

 101   TTAGGTCACTTAACAAAAAGGAAATTGGGTAGGGGTTTTTCACAGACCGC

 151   TTTCTAAGGGTAATTTTAAAATATCTGGGAAGTCCCTTCCACTGCTGTGT

 201   TCCAGAAGTGTTGGTAAACAGCCCACAAATGTCAACAGCAGAAACATACA

 251   AGCTGTCAGCTTTGCACAAGGGCCCAACACCCTGCTCATCAAGAAGCACT

 301   GTGGTTGCTGTGTTAGTAATGTGCAAAACAGGAGGCACATTTTCCCCACC

 351   TGTGTAGGTTCCAAAATATCTAGTGTTTTCATTTTTACTTGGATCAGGAA

 401   CCCAGCACTCCACTGGATAAGCATTATCCTTATCCAAAACAGCCTTGTGG

 451   TCAGTGTTCATCTGCTGACTGTCAACTGTAGCATTTTTTGGGGTTACAGT

 501   TTGAGCAGGATATTTGGTCCTGTAGTTTGCTAACACACCCTGCAGCTCCA

 551   AAGGTTCCCCACCAACAGCAAAAAAATGAAAATTTGACCCTTGAATGGGT

 601   TTTCCAGCACCATTTTCATGAGTTTTTTGTGTCCCTGAATGCAAGTTTAA

 651   CATAGCAGTTACCCCAATAACCTCAGTTTTAACAGTAACAGCTTCCCACA

 701   TCAAAATATTTCCACAGGTTAAGTCCTCATTTAAATTAGGCAAAGGAATT

 751   CTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTC

 801   ATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGT

 851   GCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATC

 901   CGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGG

 951   AAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGC

1001   GGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA
```

Figure 13

2151 TAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGC

2201 TCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTC

2251 TTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCG

2301 GGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTA

2351 CACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC

2401 CCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACA

2451 GGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAG

2501 TCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCT

2551 CGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTA

2601 CGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTT

2651 ATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATA

2701 CCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAA

2751 GCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTAT

2801 TTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCAT

2851 AGTTAAGCCAGTATACAATCAATATTGGCCATTAGCCATATTATTCATTG

2901 GTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTAT

2951 CCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCC

3001 ATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGT

3051 CATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTA

3101 AATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAAT

3151 AATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTC

3201 AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTG

3251 TATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCC

(Continued)

Figure 13 (Continued)

```
3301 CGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGC

3351 AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG

3401 CAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG

3451 TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAAC

3501 GGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGC

3551 GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAA

3601 CCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAA

3651 GACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACG

3701 CGGATTCCCCGTGCCAAGAAAGCTTGTCTAGAACCCGGGAGAGCTCCTGA

3751 GAACTTCAGGGTGAGTTTGGGGACCCTTGATTGTTCTTTCTTTTTCGCTA

3801 TTGTAAAATTCATGTTATATGGAGGGGGCAAAGTTTTCAGGGTGTTGTTT

3851 AGAATGGGAAGATGTCCCTTGTATCACCATGGACCCTCATGATAATTTTG

3901 TTTCTTTCACTTTCTACTCTGTTGACAACCATTGTCTCCTCTTATTTTCT

3951 TTTCATTTTCTGTAACTTTTTCGTTAAACTTTAGCTTGCATTTGTAACGA

4001 ATTTTTAAATTCACTTTTGTTTATTTGTCAGATTGTAAGTACTTTCTCTA

4051 ATCACTTTTTTTTCAAGGCAATCAGGGTATATTATATTGTACTTCAGCAC

4101 AGTTTTAGAGAACAATTGTTATAATTAAATGATAAGGTAGAATATTTCTG

4151 CATATAAATTCTGGCTGGCGTGGAAATATTCTTATTGGTAGAAACAACTA

4201 CATCCTGGTCATCATCCTGCCTTTCTCTTTATGGTTACAATGATATACAC

4251 TGTTTGAGATGAGGATAAAAATACTCTGAGTCCAAACCGGGCCCCTCTGCT

4301 AACCATGTTCATGCCTTCTTCTTTTTCCTACAGCTCCTGGGCAACGTGCT

4351 GGTTGTTGTGCTGTCTCATCATTTTGGCAAAGAATTCACTCCTCAGGTGC

4401 AGGCTGCCTATCAGAAGGTGGTGGCTGGTGTGGCCAATGCCCTGGCTCAC
```

                                                        (Continued)

Figure 13 (Continued)

1051 AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGAT

1101 CTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCC

1151 AATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTG

1201 TTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAAT

1251 GACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCAT

1301 GACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTG

1351 CGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCT

1401 TTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACC

1451 GGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTG

1501 CAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACT

1551 CTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGC

1601 AGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATA

1651 AATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGG

1701 CCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCA

1751 GGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCAC

1801 TGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAG

1851 ATTGATTTAAAACTTCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCT

1901 TTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACT

1951 GAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTT

2001 TTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGC

2051 GGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAA

2101 CTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCG

(Continued)

19/19

Figure 13 (Continued)

4451 AAATACCACTGAGATCTTTTTCCCTCTGCCAAAAATTATGGGGACATCAT

4501 GAAGCCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAAATTTATTTTCA

4551 TTGCAATAGTGTGTTGGAATTTTTTTGTGTCTCTCACTCGGAAGGACATAT

4601 GGGAGGGCAAATCATTTAAAACATCAGAATGAGTATTTGGTTTAGAGTTT

4651 GGCAACATATGCCATATGCTGGCTGCCATGAACAAAGGTGGCTATAAAGA

4701 GGTCATCAGTATATGAAACAGCCCCCTGCTGTCCATTCCTTATTCCATAG

4751 AAAAGCCTTGACTTGAGGTTAGATTTTTTTTTATATTTTGTTTTGTGTTAT

4801 TTTTTTCTTTAACATCCCTAAAATTTTCCTTACATGTTTTACTAGCCAGA

4851 TTTTTCCTCCTCTCCTGACTACTCCCAGTCATAGCTGTCCCTCTTCTCTG

4901 GATCC

SEQUENCE LISTING

<110> MENARINI RICERCHE S.p.A.

<120> PHARMACEUTICAL COMPOSITION, CONTAINING FRAGMENTS OF AN
      ANTIGENIC PROTEIN ENCODING DNA ENDOWED WITH ANTI-TUMOR
      EFFECT

<130> 5653MEUR

<140>
<141>

<150> MI98A002330
<151> 1998-10-30

<160> 35

<170> PatentIn Ver. 2.1

<210> 1
<211> 213
<212> DNA
<213> human

<400> 1
atgacaggtt ctggtcatgc aagctctacc ccaggtggag aaaaggagac ttcggctacc 60
cagagaagtt cagtgcccag ctctactgag aagaatgctg tgagtatgac cagcagcgta 120
ctctccagcc acagccccgg ttcaggctcc tccaccactc agggacagga tgtcactctg 180
gccccggcca cggaaccagc ttcaggttga taa   213

<210> 2
<211> 525
<212> DNA
<213> human

<400> 2
atggtgccca gctctactga gaagaatgct gtgagtatga ccagcagcgt actctccagc 60
cacagccccg gttcaggctc ctccaccact cagggacagg atgtcactct ggccccggcc 120
acggaaccag cttcaggttc agctgccacc tggggacagg atgtcacctc ggtccagtc 180
accaggccag ccctgggctc caccaccccg ccagcccacg atgtcacctc agccccggac 240
aacaagccag ccccgggaag tactgctcca ccagcacacg gtgttacctc ggctccggat 300
accaggccgg ccccaggtag taccgcccct cctgcccatg gtgtcacatc tgccccggac 360
aacaggcctg cattgggtag tacagcaccg ccagtacaca acgttactag tgcctcaggc 420
tctgctagcg gctcagcttc tactctggtg cacaacggca cctctgcgcg cgcgaccaca 480
accccagcga gcaagagcac tccattctca attcccagct gataa   525

1

<210> 3
<211> 654
<212> DNA
<213> human

<400> 3
atgggctcag cttctactct ggtgcacaac ggcacctctg ccagggctac cacaacccca 60
gccagcaaga gcactccatt ctcaattccc agccaccact ctgatactcc taccaccctt 120
gccagccata gcaccaagac tgatgccagt agcactcacc atagcacggt acctcctctc 180
acctcctcca atcacagcac ttctccccag ttgtctactg gggtctcttt cttttccctg 240
tcttttcaca tttcaaacct ccagtttaat tcctctctgg aagatcccag caccgactac 300
taccaagagc tgcagagaga catttctgaa atgttttttgc agatttataa acaaggggggt 360
tttctgggcc tctccaatat taagttcagg ccaggatctg tggtggtaca attgactctg 420
gccttccgag aaggtaccat caatgtccac gacgtggaga cacagttcaa tcagtataaa 480
acggaagcag cctctcgata taacctgacg atctcagacg tcagcgtgag tgatgtgcca 540
tttcctttct ctgcccagtc tggggctggg gtgccaggct ggggcatcgc gctgctggtg 600
ctggtctgtg ttctggttgc gctggccatt gtctatctca ttgccttgtg ataa   654

<210> 4
<211> 285
<212> DNA
<213> human

<400> 4
atgctggtgc tggtctgtgt tctggttgcg ctggccattg tctatctcat tgccttggct 60
gtctgtcagt gccgccgaaa gaactacggg cagctggaca tctttccagc ccgggatacc 120
taccatccta tgagcgagta ccccacctac cacacccatg ggcgctatgt gccccctagc 180
agtaccgatc gtagccccta tgagaaggtt tctgcaggta atggtggcag cagcctctct 240
tacacaaacc cagcagtggc agccacttct gccaacttgt gataa   285

<210> 5
<211> 1371
<212> DNA
<213> human

<400> 5
atgacaggtt ctggtcatgc aagctctacc ccaggtggag aaaaggagac ttcggctacc 60
cagagaagtt cagtgcccag ctctactgag aagaatgctg tgagtatgac cagcagcgta 120
ctctccagcc acagccccgg ttcaggctcc tccaccactc agggacagga tgtcactctg 180
gccccggcca cggaaccagc ttcaggttca gctgccacct ggggacagga tgtcacctcg 240
gtcccagtca ccaggccagc cctgggctcc accacccgc cagcccacga tgtcacctca 300
gccccggaca acagccagc ccgggaagt accgctccac cagcacacgg tgttacctcg 360
gctccggata ccaggccggc cccaggtagt accgcccctc ctgcccatgg tgtcacatct 420
gccccggaca acaggcctgc attgggtagt acagcaccgc cagtacacaa cgttactagt 480
gcctcaggct ctgctagcgg ctcagcttct actctggtgc acaacggcac ctctgcgcgc 540

```
gcgaccacaa ccccagcgag caagagcact ccattctcaa ttcccagcca ccactctgat 600
actcctacca cccttgccag ccatagcacc aagactgatg ccagtagcac tcaccatagc 660
acggtacctc ctctcacctc ctccaatcac agcacttctc cccagttgtc tactggggtc 720
tctttctttt tcctgtcttt tcacatttca aacctccagt ttaattcctc tctggaagat 780
cccagcaccg actactacca agagctgcag agagacattt ctgaaatgtt tttgcagatt 840
tataaacaag ggggtttttct gggcctctcc aatattaagt tcaggccagg atctgtggtg 900
gtacaattga ctctggcctt ccgagaaggt accatcaatg tccacgacgt ggagacacag 960
ttcaatcagt ataaaacgga agcagcctct cgatataacc tgacgatctc agacgtcagc 1020
gtgagtgatg tgccatttcc tttctctgcc cagtctgggg ctggggtgcc aggctggggc 1080
atcgcgctgc tggtgctggt ctgtgttctg gttgcgctgg ccattgtcta tctcattgcc 1140
ttggctgtct gtcagtgccg ccgaaagaac tacgggcagc tggacatctt tccagcccgg 1200
gatacctacc atccatgag cgagtacccc acctaccaca cccatggggcg ctatgtgcc 1260
cctagcagta ccgatcgtag ccctatgag aaggtttctg caggtaatg tggcagcagc 1320
ctctcttaca caaacccagc agtggcagcc acttctgcca acttgtgata a        1371
```

```
<210> 6
<211> 369
<212> DNA
<213> human
```

```
<400> 6
atgcagatct tcgtgaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
gacaccattg agaatgtcaa ggcaaagatc caagacaagg aaggcatccc tcctgaccag 120
cagaggctca tctttgcagg caagcagctg gaagatggcc gcactctttc tgactacaac 180
atccagaaag agtccaccct gcacctggtg ctccgtctca gaggtgggag gcacggtagt 240
ggtgcatggc tgttgcccgt ctcgctggtg aaaagaaaaa ccaccctggc gcccaatacg 300
caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc 360
cgaggatcc                                                     369
```

```
<210> 7
<211> 579
<212> DNA
<213> human
```

```
<400> 7
atgcagatct tcgtgaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
gacaccattg agaatgtcaa ggcaaagatc caagacaagg aaggcatccc tcctgaccag 120
cagaggctca tctttgcagg caagcagctg gaagatggcc gcactctttc tgactacaac 180
atccagaaag agtccaccct gcacctggtg ctccgtctca gaggtgggag gcacggtagt 240
ggtgcatggc tgttgcccgt ctcgctggtg aaaagaaaaa ccaccctggc gcccaatacg 300
caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc 360
cgaggatcca caggttctgg tcatgcaagc tctaccccag gtggagaaaa ggagacttcg 420
gctacccaga gaagttcagt gcccagctct actgagaaga atgctgtgag tatgaccagc 480
agcgtactct ccagccacag ccccggttca ggctcctcca ccactcaggg acaggatgtc 540
actctggccc cggccacgga accagcttca ggttgataa                     579
```

<210> 8
<211> 891
<212> DNA
<213> human

<400> 8
```
atgcagatct tcgtgaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
gacaccattg agaatgtcaa ggcaaagatc caagacaagg aaggcatccc tcctgaccag 120
cagaggctca tctttgcagg caagcagctg gaagatggcc gcactctttc tgactacaac 180
atccagaaag agtccaccct gcacctggtg ctccgtctca gaggtgggag gcacggtagt 240
ggtgcatggc tgttgcccgt ctcgctggtg aaaagaaaaa ccaccctggc gcccaatacg 300
caaaccgcct ctcccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc 360
cgaggatccg tgcccagctc tactgagaag aatgctgtga gtatgaccag cagcgtactc 420
tccagccaca gccccggttc aggctcctcc accactcagg acaggatgt cactctggcc 480
ccggccacgg aaccagcttc aggttcagct gccacctggg acaggatgt cacctcggtc 540
ccagtcacca ggccagccct gggctccacc acccccgccag cccacgatgt cacctcagcc 600
ccggacaaca agccagcccc gggaagtact gctccaccag cacacggtgt tacctcggct 660
ccggatacca ggccggcccc aggtagtacc gcccctcctg cccatggtgt cacatctgcc 720
ccggacaaca ggcctgcatt gggtagtaca gcaccgccag tacacaacgt tactagtgcc 780
tcaggctctg ctagcggctc agcttctact ctggtgcaca acggcacctc tgcgcgcgcg 840
accacaaccc cagcgagcaa gagcactcca ttctcaattc ccagctgata a       891
```

<210> 9
<211> 1020
<212> DNA
<213> human

<400> 9
```
atgcagatct tcgtgaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
gacaccattg agaatgtcaa ggcaaagatc caagacaagg aaggcatccc tcctgaccag 120
cagaggctca tctttgcagg caagcagctg gaagatggcc gcactctttc tgactacaac 180
atccagaaag agtccaccct gcacctggtg ctccgtctca gaggtgggag gcacggtagt 240
ggtgcatggc tgttgcccgt ctcgctggtg aaaagaaaaa ccaccctggc gcccaatacg 300
caaaccgcct ctcccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc 360
cgaggatccg gctcagcttc tactctggtg cacaacggca cctctgccag ggctaccaca 420
accccagcca gcaagagcac tccattctca attcccagcc accactctga tactcctacc 480
acccttgcca gccatagcac caagactgat gccagtagca ctcaccatag cacggtaccc 540
cctctcacct cctccaatca cagcacttct ccccagttgt ctactggggt ctctttcttt 600
ttcctgtctt ttcacatttc aaacctccag tttaattcct ctctggaaga tcccagcacc 660
gactactacc aagagctgca gagagacatt tctgaaatgt ttttgcagat ttataaacaa 720
gggggttttc tgggcctctc caatattaag ttcaggccag gatctctggt ggtacaattg 780
actctggcct tccgagaagg taccatcaat gtccacgacg tggagacaca gttcaatcag 840
tataaaacgg aagcagcctc tcgatataac ctgacgatct cagacgtcag cgtgagtgat 900
gtgccatttc ctttctctgc ccagtctggg gctggggtgc caggctgggg catcgcgctg 960
ctggtgctgg tctgtgttct ggttgcgctg gccattgtct atctcattgc cttgtgataa 1020
```

<210> 10
<211> 651
<212> DNA
<213> human

<400> 10

```
atgcagatct tcgtgaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
gacaccattg agaatgtcaa ggcaaagatc caagacaagg aaggcatccc tcctgaccag 120
cagaggctca tctttgcagg caagcagctg gaagatggcc gcactctttc tgactacaac 180
atccagaaag agtccaccct gcacctggtg ctccgtctca gaggtgggag gcacggtagt 240
ggtgcatggc tgttgcccgt ctcgctggtg aaaagaaaaa ccaccctggc gcccaatacg 300
caaaccgcct ctcccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc 360
cgaggatccc tggtgctggt ctgtgttctg gttgcgctgg ccattgtcta tctcattgcc 420
ttggctgtct gtcagtgccg ccgaaagaac tacggcagc tggacatctt ccagcccgg 480
gatacctacc atcctatgag cgagtacccc acctaccaca cccatgggcg ctatgtgccc 540
cctagcagta ccgatcgtag cccctatgag aaggtttctg caggtaatgg tggcagcagc 600
ctctcttaca caaacccagc agtggcagcc acttctgcca acttgtgata a         651
```

<210> 11
<211> 1737
<212> DNA
<213> human

<400> 11

```
atgcagatct tcgtgaagac cctgactggt aagaccatca ctctcgaagt ggagccgagt 60
gacaccattg agaatgtcaa ggcaaagatc caagacaagg aaggcatccc tcctgaccag 120
cagaggctca tctttgcagg caagcagctg gaagatggcc gcactctttc tgactacaac 180
atccagaaag agtccaccct gcacctggtg ctccgtctca gaggtgggag gcacggtagt 240
ggtgcatggc tgttgcccgt ctcgctggtg aaaagaaaaa ccaccctggc gcccaatacg 300
caaaccgcct ctcccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc 360
cgaggatcca caggttctgg tcatgcaagc tctaccccag gtggagaaaa ggagacttcg 420
gctacccaga gaagttcagt gcccagctct actgagaaga atgctgtgag tatgaccagc 480
agcgtactct ccagccacag ccccggttca ggctcctcca ccactcaggg acaggatgtc 540
actctggccc cggccacgga accagcttca ggttcagctg ccacctgggg acaggatgtc 600
acctcggtcc cagtcaccag gccagccctg ggctccacca cccgccagc cacgatgtc 660
acctcagccc cggacaacaa gccagccccg ggaagtaccg ctccaccagc acacggtgtt 720
acctcggctc cggataccag gccggcccca ggtagtaccg cccctcctgc ccatggtgtc 780
acatctgccc cggacaacag gcctgcattg ggtagtacag caccgccagt acacaacgtt 840
actagtgcct caggctctgc tagcggctca gcttctactc tggtgcacaa cggcacctct 900
gcgcgcgcga ccacaacccc agcgagcaag agcactccat tctcaattcc cagccaccac 960
tctgatactc ctaccaccct tgccagccat agcaccaaga ctgatgccag tagcactcac 1020
catagcacgg tacctcctct cacctcctcc aatcacagca cttctcccca gttgtctact 1080
ggggtctctt tctttttcct gtcttttcac atttcaaacc tccagtttaa ttcctctctg 1140
gaagatccca gcaccgacta ctaccaagag ctgcagagag acatttctga aatgtttttg 1200
cagatttata aacaaggggg ttttctgggc ctctccaata ttaagttcag gccaggatct 1260
gtggtggtac aattgactct ggccttccga gaaggtacca tcaatgtcca cgacgtggag 1320
```

```
acacagttca atcagtataa aacggaagca gcctctcgat ataacctgac gatctcagac 1380
gtcagcgtga gtgatgtgcc atttcctttc tctgcccagt ctggggctgg ggtgccaggc 1440
tggggcatcg cgctgctggt gctggtctgt gttctggttg cgctggccat tgtctatctc 1500
attgccttgg ctgtctgtca gtgccgccga aagaactacg ggcagctgga catctttcca 1560
gcccgggata cctaccatcc tatgagcgag tacccccacc taccacaccca tgggcgctat 1620
gtgcccccta gcagtaccga tcgtagcccc tatgagaagg tttctgcagg taatggtggc 1680
agcagcctct cttacacaaa cccagcagtg gcagccactt ctgccaactt gtgataa    1737
```

```
<210> 12
<211> 4905
<212> DNA
<213> human
```

```
<400> 12
ccaggaagct cctctgtgtc ctcataaacc ctaacctcct ctacttgaga ggacattcca 60
atcataggct gcccatccac cctctgtgtc ctcctgttaa ttaggtcact taacaaaaag 120
gaaattgggt aggggttttt cacagaccgc tttctaaggg taattttaaa atatctggga 180
agtcccttcc actgctgtgt tccagaagtg ttggtaaaca gcccacaaat gtcaacagca 240
gaaacataca agctgtcagc tttgcacaag ggcccaacac cctgctcatc aagaagcact 300
gtggttgctg tgttagtaat gtgcaaaaca ggaggcacat tttccccacc tgtgtaggtt 360
ccaaaatatc tagtgttttc atttttactt ggatcaggaa cccagcactc cactggataa 420
gcattatcct tatccaaaac agccttgtgg tcagtgttca tctgctgact gtcaactgta 480
gcattttttg gggttacagt ttgagcagga tatttggtcc tgtagtttgc taacacaccc 540
tgcagctcca aaggttcccc accaacagca aaaaaatgaa aatttgaccc ttgaatgggt 600
tttccagcac catttttcatg agttttttgt gtccctgaat gcaagtttaa catagcagtt 660
accccaataa cctcagtttt aacagtaaca gcttcccaca tcaaaatatt tccacaggtt 720
aagtcctcat ttaaattagg caaaggaatt cttgaagacg aaagggcctc gtgatacgcc 780
tattttttata ggttaatgtc atgataataa tggtttctta gacgtcaggt ggcactttttc 840
ggggaaatgt gcgcggaacc cctatttgtt tattttttcta aatacattca aatatgtatc 900
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 960
gtattcaaca tttccgtgtc gcccttattc cctttttttgc ggcattttgc cttcctgttt 1020
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 1080
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 1140
aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgtg 1200
ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 1260
agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 1320
gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag 1380
gaccgaagga gctaaccgct tttttgcaca acatgggggga tcatgtaact cgccttgatc 1440
gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 1500
cagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc 1560
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 1620
cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg 1680
gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 1740
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 1800
tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag attgatttaa 1860
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca 1920
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 1980
```

```
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaccac 2040
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa 2100
ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg tagttaggcc 2160
accacttcaa gaactctgta gcaccgccta cataccctcgc tctgctaatc ctgttaccag 2220
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 2280
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 2340
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 2400
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 2460
cgaggagct tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgccacc 2520
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 2580
ccagcaacgc ggcctttta cggttcctgg ccttttgctg gccttttgct cacatgttct 2640
ttcctgcgtt atccctgat tctgtggata accgtattac cgcctttgag tgagctgata 2700
ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc 2760
gcctgatgcg gtattttctc cttacgcatc tgtgcggtat ttcacaccgc atatggtgca 2820
ctctcagtac aatctgctct gatgccgcat agttaagcca gtatacaatc aatattggcc 2880
attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca 2940
tacgttgtat ccatatcata atatgtacat ttatattggc tcatgtccaa cattaccgcc 3000
atgttgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 3060
tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 3120
gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 3180
agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 3240
acatcaagtg tatcatatgc caagtacgcc ccctattgac gtcaatgacg gtaaatggcc 3300
cgcctggcat tatgcccagt acatgacctt atgggacttt cctacttggc agtacatcta 3360
cgtattagtc atcgctatta ccatggtgat gcggttttgg cagtacatca atgggcgtgg 3420
atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt 3480
gttttggcac caaaatcaac gggacttcc aaaatgtcgt aacaactccg ccccattgac 3540
gcaaatgggc ggtaggcgtg tacggtggga ggtctatata agcagagctc gtttagtgaa 3600
ccgtcagatc gcctggagac gccatccacg ctgttttgac ctccatagaa gacaccggga 3660
ccgatccagc ctccgcggcc gggaacggtg cattggaacg cggattcccc gtgccaagaa 3720
agcttgtcta gaacccggga gagctcctga gaacttcagg gtgagtttgg ggaccettga 3780
ttgttctttc tttttcgcta ttgtaaaatt catgttatat ggagggggca aagtttttcag 3840
ggtgttgttt agaatgggaa gatgtccctt gtatcaccat ggaccctcat gataatttttg 3900
tttctttcac tttctactct gttgacaacc attgtctcct cttattttct tttcattttc 3960
tgtaactttt tcgttaaact ttagcttgca tttgtaacga attttaaat tcacttttgt 4020
ttatttgtca gattgtaagt acttctctcta atcacttttt tttcaaggca atcagggtat 4080
attatattgt acttcagcac agttttagag aacaattgtt ataattaaat gataaggtag 4140
aatatttctg catataaatt ctggctggcg tggaaatatt cttattggta gaaacaacta 4200
catcctggtc atcatcctgc ctttctcttt atggttacaa tgatatacac tgtttgagat 4260
gaggataaaa tactctgagt ccaaaccggg cccctctgct aaccatgttc atgccttctt 4320
ctttttccta cagctcctgg gcaacgtgct ggttgttgtg ctgtctcatc attttggcaa 4380
agaattcact cctcaggtgc aggctgccta tcagaaggtg gtggctggtg tggccaatgc 4440
cctggctcac aaataccact gagatctttt tccctctgcc aaaaattatg gggacatcat 4500
gaagcccctt gagcatctga cttctggcta taaaggaaa tttattttca ttgcaatagt 4560
gtgttggaat ttttgtgtc tctcactcgg aaggacatat gggagggcaa atcatttaaa 4620
acatcagaat gagtatttgg tttagagttt ggcaacatat gccatatgct ggctgccatg 4680
aacaaaggtg gctataaaga ggtcatcagt atatgaaaca gcccctgct gtccattcct 4740
tattccatag aaaagccttg acttgaggtt agattttttt tatattttgt tttgtgttat 4800
ttttttcttt aacatcccta aaatttttcct tacatgtttt actagccaga ttttttcctcc 4860
```

tctcctgact actcccagtc atagctgtcc ctcttctctg gatcc   4905

<210> 13
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 13
gatcggatcc acaggttctg gtcatgcaag c   31

<210> 14
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 14
gatctctaga aagcttatca acctgaagct ggttccgtgg c   41

<210> 15
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 15
gatcggatcc gtgcccagct ctactgagaa gaatgc   36

<210> 16
<211> 49
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 16
gatctctaga aagcttatca gctgggaatt gagaatggag tgctcttgc   49

<210> 17
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 17
gatcggatcc ggctcagctt ctactctggt gcacaacggc   40

<210> 18
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 18
gatctctaga aagcttatca caaggcaatg agatagacaa tggcc   45

<210> 19
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 19
gatcggatcc ctggtgctgg tctgtgtcct ggttgcgc   38

<210> 20
<211> 41

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 20
gatctctaga aagcttatca caagttggca gaagtggctg c   41

<210> 21
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 21
gatctctaga atgacaggtt ctggtcatgc aagc   34

<210> 22
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 22
gatctctaga atggtgccca gctctactga gaagaatgc   39

<210> 23
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 23
gatctctaga atgggctcag cttctactct ggtgcacaac ggc   43

```
<210> 24
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 24
gatctctaga atgctggtgc tggtctgtgt tctggttgcg c                    41


<210> 25
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 25
ggcggtggag cccggggctg gcttgt                                     26


<210> 26
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 26
aacctgaagc tggttccgtg gc                                         22


<210> 27
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
```

oligonucleotide

<400> 27
gtgcccagct ctactgagaa gaatgc   26

<210> 28
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 28
gctgggaatt gagaatggag tgctcttgc   29

<210> 29
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 29
ggctcagctt ctactctggt gcacaacggc   30

<210> 30
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 30
caaggcaatg agatagacaa tggcc   25

<210> 31
<211> 27
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
     oligonucleotide

<400> 31
ctggtgctgg tctgtgttct ggttgcg   27

<210> 32
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
     oligonucleotide

<400> 32
gatctctaga atgcagatct tcgtgaagac cctgactggt   40

<210> 33
<211> 68
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
     oligonucleotide

<400> 33
tcaccagcga gacgggcaac agccatgcac cactaccgtg cctcccacct ctgagacgga 60
gcaccagg   68

<210> 34
<211> 66
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
     oligonucleotide

<400> 34
cctccgtctc agaggtggga ggcacggtag tggtgcatgg ctgttgcccg tctcgctggt 60

gaaaag   66

<210> 35
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      oligonucleotide

<400> 35
gatcggatcc tcgggaaacc tgtcgtgcca gctgc   35